Efficient and Flexible Bond Option Valuation in the Heath Jarrow and Morton Framework

Andrew Carverhill
University of Science & Technology, Kowloon, Hong Kong
visiting School of Business, Indianan University, USA

Kin Pang
The Financial Options Research Centre
Warwick Business School, The University of Warwick

February 1995

Financial Options Research Centre
Warwick Business School
University of Warwick
Coventry
CV4 7AL
Phone: 01203 524118

FORC Preprint: 95/55

Efficient and Flexible Bond Option Valuation in the Heath Jarrow and Morton Framework

Andrew Carverhill
and
Kin Pang

Abstract

The HJM bond option valuation framework is very flexible; we present an efficient numerical implementation, which uses a Monte Carlo simulation technique, with carefully chosen Martingale Variance Reduction variates. These variates make the simulation technique up to about 16 times faster, to achieve a given standard error. We also show how to ensure that the model avoids negative interest rates in this context.

EFFICIENT AND FLEXIBLE BOND OPTION VALUATION IN THE HEATH, JARROW, AND MORTON FRAMEWORK

ANDREW CARVERHILL AND KIN PANG

ANDREW CARVERHILL is professor of finance at Hong Kong University of Science and Technology in Kowloon, Hong Kong.

KIN PANG is a Ph.D. candidate at the Financial Options Research Centre at Warwick University in the United Kingdom.

Term structure models have contrasting strengths and weaknesses when applied to valuing options on the term structure. The Heath, Jarrow, and Morton [1992] model is more flexible, and takes as its inputs the initial term structure, and the volatility of the term structure. This is convenient for the user, who can directly observe the initial term structure and can estimate, or take a view on, the volatility. Moreover, the HJM model gives the option value elegantly as an expectation involving these inputs, and does not require the risk premium of the market for this. The model is thus reminiscent of the classical Black Scholes model for equity options.

The equilibrium (or factor) models, which begin with the models of Vasicek [1977] and Cox, Ingersoll, and Ross (CIR) [1985], by contrast, start from a collection of randomly evolving state variables (or factors, which usually include the short rate). These, together with a risk premium associated with each one, are taken to characterize the entire term structure at any time. Thus, the term structure and its volatility structure are outputs, and not inputs, in the equilibrium model. This can be inconvenient to the user, because these outputs may not perfectly match what is observed or desired. This can lead to the model returning values that are clearly inaccurate, because they are appropriate to the model and not what is empirically observed or desired.

At the same time, option valuation is often easier in the equilibrium models than in more general versions of the HJM model. The former require solving a diffusion equation similar to the Black-Scholes equation, while the latter most generally require a Monte Carlo simulation, or a related numerical integration technique.

We attempt to combine the advantages of the HJM model and the equilibrium models. Our approach is to retain the HJM framework and Monte Carlo procedure, and to make it more efficient by using carefully chosen martingale variance reduction (MVR) variates. The MVR technique is standard in the field of stochastic simulation, and we summarize it in the appendix. The technique has been used by Carverhill and Clewlow [1994] (see also Clewlow and Carverhill [1994]) for equity-based exotic options.[2]

Our approach contrasts with that in much of the related literature. Usually the approach is to extend the equilibrium model to be more general, as in Hull and White [1990, 1993], Fong and Vasicek [1991], Longstaff and Schwartz [1992], and Duffie and Kan [1993, 1994], or to specialize the HJM model so that the Monte Carlo technique is unnecessary, as in Ho and Lee [1986] and HJM [1990, 1992].[3] Extensions of the equilibrium model tend to be incomplete, however, and specializations of the HJM model tend to detract severely from its generality.

The model of Duffie and Kan is perhaps the closest counterpart to our approach. Duffie and Kan take an exogenously given initial term structure and volatility structure, and then derive state variables (or factors) that will provide these structures through the equilibrium model. They have to contend with the highly non-linear character of this procedure, and to get a close match they may need many factors.

We concentrate here on the European call (or put) with strike price X and maturity date s, written on a bond with coupon payments $c_1, \ldots, c_n$, at times $q_1, \ldots, q_n$ beyond time s. The time 0 value of this option is given in the HJM model by the formula

$$E_0 [\{\pm \{(c_1 P_0^{q_1} R_0^{s,q_1} + \ldots + c_n P_0^{q_n} R_0^{s,q_n}) - X P_0^s R_0^{s,s}\}\}^+] \quad (1)$$

where E_0 denotes the risk-neutral expectation at time 0, and

$$R_0^{s,q} = \exp\left\{\sum_{j=1}^{m} \int_{\rho=0}^{s} \left[v_\rho^{j,q}\, dW_\rho^j - \frac{1}{2} (v_\rho^{j,q})^2\, d\rho \right]\right\} \quad (2)$$

in which $\{v_t^{j,q}\}_{j=1,\ldots,m;\, t\leq q}$ is the volatility structure, and we take "+" for a call and "–" for a put. See HJM [1992] or Carverhill [1995b].

The volatility structure is defined by the Ito equation

$$dP_t^q / P_t^q = r_t\, dt + \sum_{j=1}^{m} v_t^{j,q}\, dW_t^j \quad (3)$$

in which P_t^q is the time t price of the pure discount bond to mature at time q; r_t is the short rate at time t; and $dW_t^1, \ldots, dW_t^m$ are the differential increments of independent standard Brownian motions.

We first show how to improve the Monte Carlo simulation to solve (1) by using martingale variance reduction variates. We implement our procedure in the context of the familiar Fong and Vasicek [1991] model.

Ensuring that the interest rates stay positive in the model is an important consideration, and not just from a purist point of view. Rogers [1994] points out that even a small probability of negative interest rates can have a large effect on prices, for long terms to maturity. In fact, in implementations of the HJM model, the volatility is often assumed to be non-random, i.e., independent of the term structure itself, although this leads directly to the possibility of negative interest rates. We show how to ensure that interest rates stay positive in our procedure, and implement it in the context of the Longstaff and Schwartz [1992] model.

I. MVR VARIATES FOR MONTE CARLO SIMULATION OF THE HJM MODEL

The Monte Carlo technique is to simulate, for many independent trials, the vector $(R_0^{s,s}, R_0^{s,q_1}, \ldots, R_0^{s,q_n})$ as given in (2), and thence the magnitude

$$\{\pm \{(c_1 P_0^{q_1} R_0^{s,q_1} + \ldots + c_n P_0^{q_n} R_0^{s,q_n}) - X P_0^s R_0^{s,s}\}\}^+ \quad (4)$$

This magnitude is the discounted payoff of the option, and each trial involves taking a sample of the m-dimensional Brownian trajectory $\{(W_t^1, \ldots, W_t^m)\}_{t\in[0,s]}$, in a suitable time discretization. The "naive" Monte Carlo estimate of the option value is then just the mean of (4) over the many trials.

In general, the MVR variates associated with a Monte Carlo simulation are a vector of martingales starting from zero, which are generated simultaneously with the trials themselves. Our choice for these variates is the vector

$$\{(R_0^{t,s} - 1, R_0^{t,q_1} - 1, \ldots, R_0^{t,q_n} - 1)\}_{t\in[0,s]} \quad (5)$$

(Note that the R_0^{t,q_i}s are exponential martingales, so long as the volatility $v_t^{j,q}$ is bounded over t, j, q — see Carverhill [1995b].)

The variance-reduced trial is then

$$\{\pm \{(c_1 P_0^{q_1} R_0^{s,q_1} + \ldots +$$

$$c_n P_0^{q_n} R_0^{s,q_n}) - X P_0^s R_0^{s,s}\}\}^+ -$$

$$\{\beta_{(n+1)} (R_0^{s,s} - 1) + \beta_1 (R_0^{s,q_1} - 1) +$$

$$\ldots + \beta_n (R_0^{s,q_n} - 1)\} \quad (6)$$

The β_is here are chosen to minimize the variance of (6) (see the appendix). The MVR technique works because (6) must have the same expectation as (4), since $E_0 [R_0^{s,q_i} - 1] = 0$ for each i.

Our choice of MVR variate is efficient because the R_0^{s,q_i}s have already been calculated for (4). Also, with these R_0^{s,q_i}s, the procedure will automatically give the correct value (without any random error), if it is applied to valuing the bond itself, which corresponds to X = 0 in the call valuation. In fact, in this case the variance is clearly reduced to zero by taking $\beta_i = c_i P_0^{q_i}$ in (6).

This variance-reduced Monte Carlo procedure can be implemented directly, when the volatility structures $\{v_t^{j,q}\}_{t\le q;j=1,\ldots,m}$ are non-random, i.e., they do not depend on the term structure itself. Also, this procedure works, provided only that the volatility structures are bounded, and it will be valuable for volatility structures for which the model cannot easily be reduced to a finite dimensional factor (i.e., equilibrium) model.[4]

A case in point would be with volatility of the shape:

$$v_t^{1,t+\tau} \equiv v^1(\tau) = (\rho - \eta) \exp\{-\alpha\tau\} + \eta \quad (7)$$

$$v_t^{2,t+\tau} \equiv v^2(\tau) = \eta/2 + \eta(1 - \exp\{-\alpha\tau\}) \quad (8)$$

which we might expect to obtain from a principal components analysis of the term structure of volatility.[5]

A simple and reasonable way to allow for stochastic volatility is simply to replace (3) by

$$dP_t^q / P_t^q = r_t\, dt + \sigma_t \sum_{j=1}^{m} v_t^{j,q}\, dW_t^j \quad (9)$$

where σ_t represents the general level of volatility, and satisfies an Ito equation

$$d\sigma_t = \xi(\sigma_t)\, dt + \eta(\sigma_t)\, dWt \quad (10)$$

where dWt is the increment of standard Brownian motion, with constant correlation coefficients, say, $dW_t.dW_t^i = \rho_i dt$, with the basic factors.

It would also be easy, but more complicated, to assign different σ_ts to each random factor. Our MVR procedure can accommodate this extra volatility factor if we replace (2) by

$$R_0^{s,q} = \exp\left\{\sum_{j=1}^{m} \int_{\rho=0}^{s} \times \left[\sigma_\rho v_\rho^{j,q}\, dW\rho^j - \frac{1}{2}\sigma_\rho^2 (v_\rho^{j,q})^2\, d\rho\right]\right\} \quad (11)$$

(See HJM [1992] and Carverhill [1995b].)

The Fong and Vasicek (FV) [1991] model can be cast in this form, as in fact can all the Gaussian models, and we present numerical results for our variance-reduced Monte Carlo procedure applied to this model. The FV model starts from the equations

$$dr_t = \alpha(\bar{r} - r_t)\, dt + \sqrt{v_t}\, dx_t \quad (12)$$

$$dv_t = (\bar{v} - v_t)\, dt + \xi\sqrt{v_t}\, dy_t \quad (13)$$

for the short rate r_t and variance $v_t \equiv \sigma_t^2$, in which dx_t and dy_t are standard Brownian motions with correlation ρ, i.e., $dx_t dy_t = \rho dt$. FV then solve this model in the form $P_t^q(r, v) = \exp\{-rD(q-t) + vF(q-t) + G(q-t)\}$ for certain functions D, F, and G, and from this we can write

EXHIBIT 1 ■ Simulation Results ■ Fong and Vasicek Model

Three-Year European Call on Six-Year Pure Discount Bond

20 Steps/Year, 50,000 Simulations				100 Steps/Year, 50,000 Simulations			
Strike	0.738147	0.745603	0.753059	Strike	0.738147	0.745603	0.753059
Moneyness	99%	100%	101%	Moneyness	99%	100%	101%
Simple MC	0.010815	0.007504	0.004889	Simple MC	0.010872	0.007599	0.004998
(std. error)	0.000058	0.000049	0.000040	(std. error)	0.000059	0.000050	0.000042
MC with MVR	0.010766	0.007463	0.004855	MC with MVR	0.010953	0.007665	0.005050
(std. error)	0.000027	0.000027	0.000026	(std. error)	0.000028	0.000028	0.000027

Three-Year European Call on Six-Year 10% Semiannual Coupon Bond

20 Steps/Year, 50,000 Simulations				100 Steps/Year, 50,000 Simulations			
Strike	0.989276	0.999269	1.009262	Strike	0.989276	0.999269	1.009262
Moneyness	99%	100%	101%	Moneyness	99%	100%	101%
Simple MC	0.014142	0.009687	0.006197	Simple MC	0.014215	0.009809	0.006338
(std. error)	0.000075	0.000063	0.000052	(std. error)	0.000076	0.000065	0.000053
MC with MVR	0.014090	0.009644	0.006164	MC with MVR	0.014359	0.009933	0.006439
(std. error)	0.000014	0.000016	0.000017	(std. error)	0.000015	0.000016	0.000017

$$dP_t^q / P_t^q = r_t\, dt - \sqrt{v_t}\, D(q - t)\, dx_t + \xi \sqrt{v_t}\, F(q - t)\, dy_t \qquad (14)$$

corresponding to (9).[6] (Also, (13) corresponds to (10), if we translate from v to $\sigma \equiv \sqrt{v}$.)

Exhibit 1 gives results for this model, with parameters $\alpha = 1.5$, $\gamma = 1.0$, $\xi = 0.1$, $\rho = 0.5$, $r_0 = 0.1$, $\bar{r} = 0.1$, $v_0 = 0.01$, and $\bar{v} = 0.01$. From these results we see that the variance reduction works better for the coupon bond than for the pure discount bond; the reduction in standard error is on the order of four, rather than two. These factors four and two indicate savings on the order of about sixteen and four, respectively, in the number of iterations required to achieve a given standard error.

The reason that the variance reduction is more effective for the coupon bond is simply that there are more control variates, because there are more coupons. Recall, however, that using these variates incurs hardly any extra cost, because they have already been calculated in the Monte Carlo procedure itself.

We have chosen to implement the FV model simply because it is familiar. The strength of our procedure is its flexibility, however; it will work for a very general volatility structure, which can be chosen or estimated directly, and when there is no analytic solution and no easy reduction of the model (such as Duffie and Kan seek, and which the Fong and Vasicek model provides) to a set of factors such as (r_t, v_t) of Equations (12) and (13).

II. ENSURING NON-NEGATIVE INTEREST RATES

Perhaps the simplest way to ensure that interest rates cannot go negative is to throw a factor $\sqrt{r_t}$ into the volatility, replacing Equation (1) by

$$dP_t^q / P_t^q = r_t\, dt + \sqrt{r_t} \sum_{j=1}^{m} v_t^{j,q}\, dW_t^j \qquad (15)$$

and keeping $v_t^{j,q}$ non-random. We could also throw in a factor σ_t, but we omit this for the sake of clarity. Then, following HJM [1992], we have

$$df_t^q = r_t\, \alpha_t^q\, dt + \sqrt{r_t} \sum_{j=1}^{m} \sigma_t^{j,q}\, dW_t^j \qquad (16)$$

where

$$\sigma_t^{j,q} = -\frac{\partial}{\partial q} v_t^{j,q}, \quad \alpha_t^q = -\sum_{j=1}^{m} v_t^{j,q}\, \sigma_t^{j,q} \qquad (17)$$

so

$$r_t\ (\equiv f_t^t) = f_0^t + \int_{\rho=0}^t r_\rho\, \alpha_\rho^t\, dt + \sum_{j=1}^m \int_{\rho=0}^t \sqrt{r_t}\, \sigma_\rho^{j,t}\, dW_\rho^j \tag{18}$$

Also we have

$$R_0^{s,q} = \exp\left\{\sum_{j=1}^m \int_{\rho=0}^s \left[\sqrt{r_\rho}\, v_\rho^{j,q} dW_\rho^j - \frac{1}{2} r_\rho (v_\rho^{j,q})^2 d\rho\right]\right\} \tag{19}$$

The idea is now to simulate the $R_0^{t,q}$'s and r_t simultaneously as t evolves over [0, s], using (18) and (19), and continue the MVR procedure.

From (18) it is impossible for r_t to go negative, as long as the initial forward rates $\{f_0^q\}_{q\geq 0}$ are never negative. Also, if r_t is never negative, then neither is f_t^q for any t, q; this can be seen since the RHS always decreases in the equation

$$\exp\{-\textstyle\int_{\tau=t}^q f_t^\tau d\tau\}\ (\equiv P_t^q) = E[\exp\{-\textstyle\int_{\tau=t}^q r_\tau d\tau\}] \tag{20}$$

The Longstaff and Schwartz model fits indirectly into this framework, and we present numerical results for this model. Again, we have chosen to implement our procedure for this model simply because it is familiar, but the strength of our procedure is its generality. The formulation above, using the square root of r_t, will work as long as the volatility structures $\{v_t^{j,q}\}_{t\leq q}$ in (15) are all bounded.[7]

The Longstaff and Schwartz model starts from the equations

$$dx_t = (\gamma - \delta x_t)\, dt + \sqrt{x_t}\, dW_t^0 \tag{21}$$

$$dy_t = (\eta - \varepsilon y_t)\, dt + \sqrt{y_t}\, dW_t^1 \tag{22}$$

where (x_t, y_t) are related linearly to (r_t, v_t) via the equations $r = \alpha x + \beta y$ and $v = \alpha^2 x + \beta^2 y$, and W_t^0, W_t^1 are orthogonal Brownian motions. Longstaff and Schwartz obtain functions A, B, C, and D such that

$$P_t^q\ (r, v) = A^{2\gamma}(q-t)\, B^{2\eta}(q-t) \times \exp\{\kappa(q-t) + C(q-t)\, r + D(q-t)\, v\}$$

in which $\kappa = \gamma(\delta + \phi) + \eta(v + \varphi)$, $\phi = \sqrt{2\alpha + \delta^2}$, $\varphi = \sqrt{2\beta + v^2}$, and v is a constant.

From this we see that

$$dP_t^q / P_t^q = r_t dt + \sqrt{x}\, \alpha \times [C(q-t) + \alpha[D(q-t)]\, dW_t^0 + \sqrt{y}\, \beta [C(q-t) + \beta D(q-t)]\, dW_t^1 \tag{23}$$

With parameters $\alpha = 0.05$, $\beta = 0.06$, $\gamma = 0.8$, $\delta = 0.8$, $\varepsilon = 0.5$, $\eta = 0.9$, and $v = 0.5$, the results are given in Exhibit 2. Our conclusions for the LS implementation are essentially the same as for the FV implementation. The variance reduction is better for the coupon bond, and requires fewer iterations for a given standard error, by a factor of about nine.

III. SUMMARY AND COMMENTARY

Our implementation of the HJM model for valuing options on the term structure is efficient and flexible, and can avoid the possibility of negative interest rates. To ensure flexibility, our implementation is basically Monte Carlo simulation; we have made this more efficient by using martingale variance reduction variates.

The MVR technique is well-established in the field of stochastic simulation, and our contribution has been to make an appropriate choice of MVR variates. When valuing a coupon bond option, our MVR variates are associated with the pure discount bonds that make up the coupons. These must be calculated when simulating the option payoff itself, so using these MVR variates essentially imposes no extra cost in calculation. To ensure that interest rates never become negative, we insert the factor $\sqrt{r_t}$ in a way that mimics the CIR model.

When we have cast the models of Fong and Vasicek [1991] and Longstaff and Schwartz [1992] in our framework, and apply our variance-reduced Monte Carlo procedure, we get more variance reduction when there are more coupons. The reason is that there are then more MVR variates. For a three-year option on a six-year bond that pays semiannual coupons (so that

EXHIBIT 2 ■ Simulation Results ■ Longstaff and Schwartz Model

Three-Year European Call on Six-Year Pure Discount Bond							
20 Steps/Year, 50,000 Simulations				100 Steps/Year, 50,000 Simulations			
Strike	0.646075	0.652601	0.659127	Strike	0.646075	0.652601	0.659127
Moneyness	99%	100%	101%	Moneyness	99%	100%	101%
Simple MC	0.020339	0.017744	0.015321	Simple MC	0.020581	0.017977	0.015545
(std. error)	0.000113	0.000104	0.000096	(std. error)	0.000113	0.000105	0.000097
MC with MVR	0.020390	0.017791	0.015363	MC with MVR	0.020552	0.017951	0.015522
(std. error)	0.000057	0.000056	0.000055	(std. error)	0.000057	0.000056	0.000055
Three-Year European Call on Six-Year 10% Semiannual Coupon Bond							
20 Steps/Year, 50,000 Simulations				100 Steps/Year, 50,000 Simulations			
Strike	0.879675	0.888560	0.897446	Strike	0.879675	0.888560	0.897446
Moneyness	99%	100%	101%	Moneyness	99%	100%	101%
Simple MC	0.025770	0.022224	0.018932	Simple MC	0.026076	0.025518	0.019213
(std. error)	0.000142	0.000131	0.000120	(std. error)	0.000143	0.000132	0.000120
MC with MVR	0.025858	0.022305	0.019005	MC with MVR	0.026056	0.025501	0.019199
(std. error)	0.000046	0.000047	0.000047	(std. error)	0.000047	0.000048	0.000048

there are seven MVR variates, corresponding to the number of coupons plus the option payment itself), the standard error for a given number of simulations is reduced by a factor of about four. This translates to a reduction in the number of simulations by a factor of about sixteen to achieve a given standard error.

The motivation for applying our procedure to these models is simply that they are familiar. The strength of our procedure is that it will work very generally — we require only some regularity conditions on the volatility of the term structure.

Is our HJM implementation preferable to its rivals, the equilibrium (factor) models? This depends on the priorities of the user. If flexibility and accuracy in fitting market data are needed, our approach is valuable. Factor models such as Duffie and Kan [1993] can achieve this flexibility by including more factors, although this may be difficult because of the basically non-linear character of this procedure. The time taken to solve the factor model will grow geometrically with the number of factors, while the time taken by the Monte Carlo procedure will grow less dramatically as the number of factors increases.

APPENDIX ■ The Martingale Variance Reduction Technique

In its "naive" form, the Monte Carlo technique may be expressed as: Suppose we want to estimate the mean, μ, of a probability distribution, and we are able to generate independent random samples (trials) $\{y_1, y_2, \ldots\}$ from the distribution. Then the simple Monte Carlo estimator of μ, based on M samples, is just the sample mean

$$\hat{\mu} = \frac{1}{M}\sum_{i=1}^{M} y_i \qquad \text{(A-1)}$$

If we denote the variance of each y_i (and hence of the distribution) by Var (y), then by the central limit theorem, the variance of $\hat{\mu}$ (i.e., its standard error squared, as an estimator of μ) is given by Var (y)/M. Also, if we can achieve a certain variance with, say, M samples, to reduce this by a factor of, say, X, will require us to generate XM samples.

Now, suppose that in parallel with generating $\{y_1, y_2, \ldots\}$ we also generate the antithetic variates $\{y_1^*, y_2^*, \ldots\}$, which are also independent and of the same distribution as the y_is, but negatively correlated with them. Then the variates $\bar{y}_i \equiv 1/2\,(y_i + y_i^*)$ will again have mean μ, but they will have variance $\mathrm{Var}(\bar{y}) \equiv 1/2\,\mathrm{Var}(y)\,[1 + \mathrm{Corr}(y, y^*)]$ $[\leq 1/2\,\mathrm{Var}(y)]$. Moreover, the sample mean

$$\hat{\hat{\mu}} = \frac{1}{M}\sum_{i=1}^{M} \bar{y}_i \qquad \text{(A-2)}$$

will have variance less than half that of $\hat{\mu}$. If the pair y_i, y_i^* is less than twice as expensive to generate than y_i, calculating $\hat{\hat{\mu}}$ will be a more efficient way to estimate μ than calculating $\hat{\mu}$. This is the antithetic variate technique.[8]

The martingale variance reduction technique is to generate simultaneously with the samples $\{y_1, y_2, \ldots\}$ the i.i.d. variates $\{x_1, x_2, \ldots\}$, where each x_i is an m-vector $(x_i^1, \ldots, x_i^m)^T$ (the symbol T denoting "transpose"), whose increments have zero mean, i.e., they are martingales. Thus, the (m + 1)-vectors (y_i, x_i) are independent for different values of i, but their components are correlated among themselves in a constant way. Then for any m-vector ξ, the quantity

$$\hat{\mu}_c = \frac{1}{M}\sum_{i=1}^{M}\left(y_i - \sum_{j=1}^{m} x_i^j \xi_j\right) \quad \text{(A-3)}$$

will also be an estimator of μ, and the variance of this is minimized if we choose ξ so that

$$\Sigma_{xx}\,\xi = \Sigma_{xy} \quad \text{(A-4)}$$

where Σ_{xx} is the covariance matrix of x, and Σ_{xy} is the m-vector of covariances between x and y.

If we know the matrix Σ_{xx} and vector Σ_{xy}, then Equations (A-3) and (A-4) let us use the control variates x_i to reduce the variance in our simple Monte Carlo estimation (A-2) of μ. If we do not know Σ_{xx} and Σ_{xy}, the procedure to obtain the variance-reduced estimator $\hat{\mu}_c$ is: Find $\hat{\beta} \equiv (\hat{\beta}_0, \ldots, \hat{\beta}_m)^T$ to fit, in a least squares sense, the system of equations

$$X_i\,\hat{\beta} = y_i, \qquad i = 1, \ldots, M \quad \text{(A-5)}$$

where $X_i = (1, x_i^1, \ldots, x_i^m)$. Then $(\hat{\beta}_1, \ldots, \hat{\beta}_m)$ is the choice of $(\xi_1, \ldots, \xi_m)$ that will minimize the variance of the estimator (A-3), and $\hat{\beta}_0$ is the corresponding variance-reduced estimator $\hat{\mu}_c$ of μ.

In fact, $\hat{\beta}$ can be calculated simply as

$$\hat{\beta} = (X^T X)^{-1} X^T y \quad \text{(A-6)}$$

provided none of the control variates is redundant. If some of the control variates are redundant, or highly collinear, then X^TX will be singular, or nearly singular.[9] In this case, however, because X^TX is symmetric, we can work with its pseudo-inverse, which is obtained by projecting along the directions given by the eigenvectors with zero eigenvalue.

The m × m matrix X^TX and m-vector X^Ty can be calculated conveniently as the simulation proceeds, using the equations

$$(X_{n+1}^T X_{n+1})_{j,k} = (X_n^T X_n)_{j,k} + x_j^{n+1} x_k^{n+1} \quad \text{(A-7)}$$

$$(X_{n+1}^T y_{n+1})_j = (X_n^T y_n)_j + x_j^{n+1} y_{n+1} \quad \text{(A-8)}$$

where $x_0^{n+1} = 1$

ENDNOTES

[1]Technically, the HJM model is not separate from the equilibrium models, but rather includes them as special cases. The HJM model takes the initial term structure and volatility structures, but does not attempt to explain them in terms of state variables.

[2]In these articles the approach to the design of the MVR variates is different; the design is based on the delta-hedge.

[3]This is a retrospective interpretation, since Ho and Lee [1986] predates HJM [1992]. The most common implementations of the HJM model use the volatility of the Vasicek [1977] model or the Cox, Ingersoll, and Ross [1985] model. See HJM [1990, 1992]. Carverhill [1994] shows that if the volatility is time-homogeneous and non-random, and if the short rate is Markovian (which is necessary for compatibility with a one-factor model), then the volatility must be as in Vasicek. Rogers [1993] makes the point that practicable implementations of the HJM model often push the model into a formulation that would more conveniently start from a factor model.

[4]See the Appendix to Carverhill [1995b].

[5]The significant aspect of this volatility structure is that (8) represents a tilting of the term structure. Such a tilt seems to be difficult to incorporate into a factor structure in a time-homogeneous way.

[6]FV obtain the functions D, F, and G from a series of transformations. Selby and Strickland [1993] obtain them as series expansions. To implement our procedure, we must transform to orthogonal Brownian motions.

[7]See the Appendix to Carverhill [1995b]. Note, however, that we cannot replace $\sqrt{r_t}$ with, say, r_t in (15), because we would not then be able to ensure that r_t, given as a solution corresponding to (18), and P_t^q itself, is well-defined and non-explosive.

[8]We have implemented this technique in the HJM framework, but we do not report the results here, because the variance reduction is quite feeble.

[9]This tends to be the case when applying this technique, because the bonds corresponding to the coupons are highly correlated.

REFERENCES

Carverhill, A.P. "A Note on the Models of Hull and White for Pricing Options on the Term Structure." *Journal of Fixed*

Income, September 1995a, pp. 89-96.

——. "A Simplified Exposition of the Heath, Jarrow and Morton Model." *Stochastics*, Vol. 53 (1995b), pp. 227-240.

——. "When is the Short Rate Markovian?" *Mathematical Finance*, Vol. 4, Number 4 (1994), pp. 305-312.

Carverhill, A.P., and L.C. Clewlow. "Computing Option Values: A Fast Monte Carlo Method ('Quicker on the Curves')." *Risk*, May 1994.

Clewlow, L.C., and A.P. Carverhill. "On the Simulation of Contingent Claims." *Journal of Derivatives*, Winter 1994, pp. 66-74.

Cox, J.C., J.E. Ingersoll, and S.A. Ross. "A Theory of the Term Structure of Interest Rates." *Econometrica*, 53 (1985), pp. 385-407.

Duffie, D., and R. Kan. "Multi-Factor Term Structure Models." *Phil. Transactions of the Royal Society of London Actuaries*, 347 (1994), pp. 577-586.

——. "A Yield Factor of Interest Rates." Working Paper, Graduate School of Business, Stanford University, 1993.

Fong, H.G., and O. Vasicek. "Interest Rate Volatility as a Stochastic Factor." Working Paper, Gifford Fong Associates, 1991.

Heath, D., R. Jarrow, and A. Morton. "Bond Pricing and the Term Structure of Interest Rates: A Discrete Time Approximation." *Journal of Financial and Quantitative Analysis*, Vol. 25, No. 4 (1990), pp. 419-440.

——. "Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation." *Econometrica*, Vol. 60, No. 1 (1992), pp. 77-105.

Ho, T.S.Y., and S.B. Lee. "Term Structure Movements and Pricing Interest Rate Contingent Claims." *Journal of Finance*, XLI (1986), pp. 1011-1029.

Hull, J., and A. White. "One Factor Interest Rate Models and the Valuation of Interest Rate Derivative Securities." *Journal of Financial and Quantitative Analysis*, Vol. 28, No. 2 (1993), pp. 235-254.

——. "Pricing Interest Rate Derivative Securities." *Review of Financial Studies*, Vol. 3, Number 4 (1990).

Longstaff, F.A., and E.S. Schwartz. "Interest Rate Volatility and the Term Structure: A Two-Factor General Equilibrium Model." *Journal of Finance*, September 1992.

Rogers, L.C.G. "Which Model of the Term-Structure of Interest Rates Should One Use?" Proceedings of the IMA Workshop on Mathematical Finance, 1993, presented at the Annual Conference of the Financial Options Research Centre, Warwick, U.K., 1994.

Selby, M.J.P., and C. Strickland. "Computing the Fong and Vasicek Pure Discount Bond Pricing Formula." Working Paper 93/42, Financial Options Research Centre, University of Warwick, U.K., 1993.

Vasicek, O. "An Equilibrium Characterization of the Term Structure." *Journal of Financial Economics*, 5 (1977), pp. 177-188.